A Biologically Based Model for the Analysis of Premalignant Foci of Arbitrary Shape

ANUP DEWANJI
Indian Statistical Institute, Computer Science Unit, Calcutta, 700035, India

AND

E. GEORG LUEBECK AND SURESH H. MOOLGAVKAR
Fred Hutchinson Cancer Research Center, Public Health Sciences Division, Seattle, Washington 98104

Received 16 March 1995; revised 27 September 1995

ABSTRACT

In many animal carcinogenesis experiments, quantitative data on putative premalignant foci are now routinely collected. Moolgavkar et al. [*Carcinogenesis* 11:1271 (1990)] considered the analysis of such data from a rat hepatocarcinogenesis experiment within the framework of a two-stage model for carcinogenesis using the assumption that the premalignant clones were spherical. This assumption seems questionable in many organs, including the liver. In this paper, it is relaxed and arbitrary shapes are allowed for the clones. The proposed method is illustrated by reanalysis of the data considered in the earlier paper. The new analysis yields parameter estimates that are more plausible biologically than those of the original analysis.

INTRODUCTION

Many experimental carcinogenesis animal model systems are characterized by the appearance of well-defined lesions that are believed to represent an early stage in the carcinogenic process. The oldest example of such an experimental model system is provided by the mouse skin initiation–promotion system, in which premalignant lesions, called papillomas, appear. Another well-studied example is provided by the rodent hepatocarcinogenesis system, which is characterized by the appearance of clones of cells exhibiting specific enzyme alterations [1, 2] that are readily recognized under the microscope after appropriate staining. The quantitative analysis of altered hepatic foci presents a unique challenge because the observations are made on two-dimensional sections of the liver. Three-dimensional reconstruction of the foci

requires the use of stereological techniques. In the current biological literature these reconstructions are usually made by using procedures due to Fullman and Saltykov (see, e.g., [3]). These procedures involve the "binning" of foci into predetermined size classes and have the unpleasant property of yielding negative numbers in some of the bins if the number of observed transections is not large enough. Nychka et al. [4] have used smoothing techniques to estimate the size distribution of foci from the size distribution of two-dimensional sections of foci.

In previous papers, we developed expressions for the distributions of the number and sizes of intermediate foci based on a biological model for initiation and promotion [5, 6]. This model, in turn, is part of the two-mutation clonal expansion model for carcinogenesis, which has been shown to be consistent with a number of epidemiologic and experimental data sets [7–9]. Statistical methods for analyses of altered hepatic foci using these expressions were presented in [10, 11]. A crucial step in the analyses was the conversion of the model-derived three-dimensional expressions for the number and sizes of foci into corresponding two-dimensional expressions for the number of transections and their areas. This method shares one important feature with the methods developed by Saltykov and Nychka: it assumes that the foci are perfectly spherical and uses a transformation due to Wicksell [12].

The assumption that foci are perfectly spherical, which has hitherto been used for analyses of altered hepatic foci, is a strong assumption. It precludes the use of these methods for the analysis of foci in other organs, such as the pancreas and the kidney, in which the foci are far from spherical. In rodent liver also, many foci appear not to be spherical. In this paper we develop the stereological methods required to fit the parametric expressions for the distributions of the number and sizes of foci derived from the two-mutation clonal expansion model to two-dimensional data without making any assumptions regarding the shape of foci. We make essential use of a fundamental result in stereology that asserts that the area fraction equals the volume fraction (see [13, p. 282]). In our context, this result states that the volume fraction occupied by altered foci is unbiasedly estimated by the observed area fraction in two-dimensional sections, whatever the shape of the foci. The intuitive idea of the method is the following. The volume fraction can be computed from the expressions derived from the two-mutation model. The volume fraction can also be estimated from the observed area fraction. By equating these two expressions, the parameters of the model can be estimated. We derive likelihood methods for implementing this basic idea.

The two-mutation clonal expansion model incorporates two important features: (1) transition of target stem cells into cancer cells via an

intermediate stage in two rate-limiting, irreversible, and hereditary (at the cell level) steps; and (2) growth and differentiation of normal target and intermediate (initiated or premalignant) cells. Suppose that $X(s)$ is the number of normal susceptible cells at age s and $\nu(s)$ is the rate of initiation of a normal cell. Once an initiated cell is created, it divides into two initiated cells, dies or differentiates, or divides into one initiated cell and one malignant cell (having sustained the second critical event for malignant transformation) with rates of $\alpha(s)$, $\beta(s)$, and $\mu(s)$, respectively. Since we are not concerned with malignant conversion of initiated cells in this paper, only α and β are relevant. Note that all parameters may depend upon the dose of the chemical agent under investigation.

The required distributional results are derived in the next section. Following that we present an analysis of data obtained from an experiment in which rats were administered NNM at various concentrations in their drinking water.

DISTRIBUTIONAL RESULTS

In order to present a motivated approach to this problem, we first consider a single clone and its corresponding focus in two dimensions and develop results on the distribution of the corresponding area fraction. Then we consider the distribution of the number of foci in two dimensions, and in the final part of this section we combine the results to obtain the results for analyzing data on the number of foci and their area fractions.

DISTRIBUTION OF AREA FRACTION

The distribution of the (random, but unobserved) volume fraction of a single clone can be obtained, at least in theory, from the results in [5] or [6]. But it is not possible to obtain the distribution of the (observed) area fraction from that of the volume fraction without any assumptions on shape. However, at least the first and second moments of the (observed) area fraction can be obtained without any such assumptions (see [13, pp. 174-178]). Let V_f and A_f denote the (unobserved) volume fraction and (observed) area fraction, respectively. Also, let V_{ic} and V_T denote the volumes of an initiated cell and of the tissue block from which the section was cut, respectively. Then, under the assumption of stationarity and isotropy for the formation of the clones, we have for an observed area fraction at time t,

$$E[A_f] = E[V_f] = E[W_1(t)] \times \frac{V_{ic}}{V_T} = f, \quad \text{say}, \tag{1}$$

where $W_j(t)$ denotes the number of cells in a single clone at time t. Also, under what Stoyan et al. [13, p. 184] call the exponential covariance assumption,

$$\mathrm{Var}[A_f] \approx \frac{\sigma^2}{A} f(1-f), \qquad \sigma^2 > 0, \tag{2}$$

where A is the total area of the cross section. These two results, (1) and (2), suggest a least squares analysis. However, in order to incorporate the concept of a threshold size for A_f below which the two-dimensional section of the focus cannot be detected, some distributional results on A_f are needed. Once these are derived, the likelihood method is readily applied.

The relationship between area fraction and the corresponding volume fraction can be established by writing the distribution of area fraction A_f as

$$h(a_f) = \int_0^1 p(a_f|v_f) g(v_f)\, dv_f, \tag{3}$$

where $h(a_f)$ denotes the density of area fraction A_f, $p(a_f|v_f)$ the conditional density of A_f given volume fraction $V_f = v_f$, and $g(v_f)$ the density of volume fraction V_f. Here lowercase letters are used to mean a particular value of the corresponding random variable. Since, for $V_f = 0$, the conditional distribution $p(a_f|f_f)$ is degenerate, we consider only $V_f > 0$ (i.e., nonextinct clones), and $g(v_f)$ is taken as the conditional density of V_f given $V_f > 0$. The reason will be clearer when we consider the distribution for the number of nonextinct foci in the next subsection. Note that the density $g(v_f)$ of V_f can be obtained from the two-stage model (as described later). Under the assumption that clones are spherical, the conditional density $p(a_f|v_f)$ can be written down using Wicksell's formula [12], in terms of the radii corresponding to a_f and v_f. To get around the assumption that clones are spherical, we proceed as follows.

Note that, given volume fraction $V_f = v_f$, the area fraction A_f is the proportion of area covered by the clone in a random test plane (or cross section). Thus, the area fraction is zero if the cross section does not intersect the clone [12, 14]. We assume that this happens (that is, $A_f = 0$) with probability $1 - v_f^{1/3}$; that is, the random cross section intersects the clone with probability $v_f^{1/3}$. Given that there is an intersection ($A_f > 0$), we assume that the conditional distribution of A_f is given by

$$A_f|V_f = v_f \sim N\left(v_f^{2/3}, \frac{\sigma^2}{A} v_f(1-v_f)\right), \tag{4}$$

where $\sigma^2 > 0$ and A is as in (2). The normality assumption is supported by the asymptotic results in [15, 16]. The expression for the variance in (4) is supported by (2) and also by the fact that near the two extremes of v_f (0 and 1), the variance will be smaller. Then the conditional distribution $P(a_f|v_f)$ of A_f given $V_f = v_f$ is

$$p(a_f|v_f) = (1 - v_f^{1/3}) I_{\{a_f = 0\}} + v_f^{1/3} \phi\left(a_f; v_f^{2/3}, \frac{\sigma^2}{A} v_f(1 - v_f)\right) I_{\{a_f > 0\}}, \quad (5)$$

where $I_{\{\cdot\}}$ denotes the indicator function and $\phi(\cdot; \mu, \sigma^2)$ the normal density with mean μ and variance σ^2.

There is the problem of detectability so that a section of a focus with area less than a threshold area (δ, say) is not observable. Also, for reasons discussed in [10], some sections with area larger than an upper bound (Δ, say) are not considered. To account for this, we need to calculate the probability that a focus is detectable, that is, the corresponding area fraction lies between δ/A and Δ/A. Let $H(\cdot)$ denote the distribution function of A_f. Then, using (5) and (3) and appropriate normal probabilities, we obtain this probability of observing a nonextinct focus as

$$[H(\Delta/A) - H(\delta/A)] = \int_0^1 v_f^{1/3} \left[\Phi\left(\frac{\Delta}{A}; v_f^{2/3}, \frac{\sigma^2}{A} v_f(1 - v_f)\right) - \Phi\left(\frac{\delta}{A}; v_f^{2/3}, \frac{\sigma^2}{A} v_f(1 - v_f)\right)\right] g(v_f)\, dv_f, \quad (6)$$

where $\Phi(\cdot; \mu, \sigma^2)$ is the distribution function corresponding to the normal density $\phi(\cdot; \mu, \sigma^2)$. The density of an observed area fraction A_f can now be obtained as

$$\frac{\int_0^1 v_f^{1/3} \phi\left(a_f; v_f^{2/3}, (\sigma^2/A) v_f(1 - v_f)\right) g(v_f)\, dv_f}{[H(\Delta/A) - H(\delta/A)]} \quad (7)$$

for $\delta/A \leqslant a_f \leqslant \Delta/A$. Note that any least squares type of method requiring only the first two moments of A_f cannot adjust for such truncation constraints in the observations.

To find $g(v_f)$, note that W_1, the size (number of cells) of a single initiated clone, is related to V_f by

$$W_1 = W_1(V_f) = \left[\frac{V_T}{V_{ic}} \times V_f\right]. \quad (8)$$

Then $g(v_t)$ can be written approximately as

$$g(v_t) = \frac{V_T}{V_{ic}} p_{W_1}(w), \tag{9}$$

where $p_{W_1}(w)$ denotes the conditional probability mass function (p.m.f.) of W_1 given $W_1 > 0$ (nonextinct clone).

Let Ψ^* and Ψ denote the probability generating functions (p.g.f.'s) of the sizes of a nonextinct clone and a clone (extinct or not), respectively, and $p(t,s)$ the probability of extinction of a clone at time t originating at time s. The p.g.f's Ψ and Ψ^* are given by

$$\Psi(t,s;u) = 1 + \frac{u-1}{g(t,s) + G(t,s)(1-u)},$$

where $g(t,s) = \exp[-\int_s^t (\alpha(u) - \beta(u))\,du]$ and $G(t,s) = \int_s^t \alpha(u) g(u,s)\,du$, and

$$\Psi^* = \frac{\Psi - p(t,s)}{1 - p(t,s)}$$

(see [7, 17]). Note that the probability of extinction $p(t,s)$ can be obtained from this expression for Ψ as $1 - [g(t,s) + G(t,s)]^{-1}$. Using the expression for Ψ^*, one can, at least in theory, obtain the p.m.f. $p_{W_1}(w)$ and hence $g(v_t)$ from (9). However, when the parameters of the two-stage model, namely νX, α, and β, are independent of time, the results simplify. One can check that, in that case, the p.m.f. for the size of a single nonextinct clone at time t is [10]

$$P[W_1(t) = m] = -\frac{1}{\log(1-\theta)} \left(\frac{\theta^m}{m}\right), \qquad m = 1,2,\ldots;\ 0 < \theta < 1, \tag{10}$$

where

$$\theta = \frac{\alpha}{\beta} p(t)$$

with

$$p(t) = p(t,0) = \frac{\beta - \beta \exp[-(\alpha - \beta)t]}{\alpha - \beta \exp[-(\alpha - \beta)t]}.$$

This is a logarithmic series distribution with p.g.f.

$$\frac{\log(1-\theta u)}{\log(1-\theta)}.$$

DISTRIBUTION OF NUMBER OF FOCI

We know that the distribution of the number of nonextinct clones at time t, $N(t)$, is Poisson $(\Lambda(t))$, where [5]

$$\Lambda(t) = \int_0^t \nu(s) X(s) \left[1 - p(t,s)\right] ds.$$

A fraction of these $N(t)$ clones, say $N_2(t)$, are cut by the random cross section, and the corresponding transections have area fractions lying between δ/A and Δ/A with probability $H(\Delta/A) - H(\delta/A)$. Thus, we have

$$N_2(t) \sim \text{Poisson}(\eta_2(t)),$$

where

$$\eta_2(t) \quad \Lambda(t)\left[H(\Delta/A) - H(\delta/A)\right]. \tag{11}$$

See [10] for a similar result. Note that, for time-independent parameters (νX, α, and β), the Poisson parameter $\Lambda(t)$ of $N(t)$ simplifies:

$$\Lambda(t) = \nu X \frac{1}{\alpha}\left[\log\left(\frac{\alpha \exp\left[(\alpha - \beta)t\right] - \beta}{\alpha - \beta}\right)\right]. \tag{12}$$

NUMBER OF FOCI AND THEIR AREA FRACTIONS

In addition to V_T, the data for each animal should ideally consist of t, d, a, n_2, and (if $n_2 > 0$), $a_1, \ldots, a_{n_2}$, where t denotes the time of observation, d the corresponding dose level of the agent under consideration, a the value of A (area of cross section), n_2 the observed value of $N_2(t)$ (number of observed transections of foci), and $a_1, \ldots, a_{n_2}$ the corresponding areas of the transections. For each i, a_i/a is the observed area fraction of the ith transection. Here V_{ic}, V_T, t, and d are assumed known and nonrandom, the value of a is random but known, and $(n_2, a_1, \ldots, a_{n_2})$ are the random quantities of interest. In the following, we develop the likelihood function based on the conditional distribution of $(n_2, a_1/a, \ldots, a_{n_2}/a)$ given $A = a$.

We developed in the preceding subsection that $N_2(t)$ follows a Poisson distribution with parameter $\eta_2(t)$ [see (11)]. Note that, given $N_2(t) = n_2$, $a_1/a, \ldots, a_{n_2}/a$ are independent and identically distributed with density (7). Thus, the likelihood contribution from an observation with n_2 foci is a product of $n_2 + 1$ terms; the first term is the Poisson probability of n_2 foci with parameter $\eta_2(t)$, and then each of the n_2 foci contributes a term of the form (7). The total likelihood is obtained by taking the product of contributions from all observations.

AN EXAMPLE

We consider the *N*-nitrosomorpholine (NNM) data analyzed by Moolgavkar et al. [10]. Experimental details can be found in that paper or in [2]. Briefly, a number of rats were administered NNM at various concentrations (0, 0.1, 1, 5, 10, 20, and 40 ppm) in the drinking water. The rats were killed at different ages, and cross sections from their livers were examined for ATPase-deficient foci. Following [10], we take the lower detection limit δ to be the area of a circle with radius 60 μm and the upper truncation limit Δ to be the area of a circle with radius 500 μm. The data then consist of 1654 transections in 162 rats. The parameter estimates are rather insensitive to changes in the detection and truncation limits.

Since, in this experiment, exposure to a single agent begins at a fixed age and continues at the same level until death or until the animal is killed, the parameters of the model are assumed to be dose-dependent but constant in time. It is convenient to consider the mean number of cells per unit volume of the tissue, X_u say, instead of $X(s)$ [10]. Then νX_u may be interpreted as the initiation rate per unit volume and is related to νX of (12) by $\nu X = \nu X_u \times V_T$. It should be noted that the parameter ν is not individually estimable as it always appears with X_u, and so we treat νX_u as a single parameter. The parameters of the model, namely νX_u (the initiation rate per unit volume) and α and β (the division and death or differentiation rates for the initiated cells, respectively), are reparameterized to νX_u, $\alpha - \beta$, and β/α as this gives better numerical stability. The first two parameters are made dose (d) dependent by $\nu X_u(d) = c_1 + c_2 d$ and $(\alpha - \beta)(d) = b_1 + b_2 d$. Likelihood ratio tests showed no evidence of dose dependence in the parameter β/α, as noted also in [10]. Thus, we have a total of six parameters, namely c_1, c_2, b_1, b_2, β/α, and σ^2. These parameters are estimated using the method of maximum likelihood, where the likelihood function for the observed data is constructed using the results of the previous section.

In this example, we have three cross sections from the liver from each rat. Unfortunately, the volumes of the tissue blocks, V_T, were not measured. For the sake of this illustration, we take V_T to be same for all rats, having the approximate value of 1.43 cm^3. This value is obtained by taking the average over all sections of $A_i^{3/2}$, where A_i is the area of the cross section for the ith observation. As in [10], we assume that V_{ic}, the volume of an initiated cell, is the same as the volume of a sphere with radius 12 μm. The effect of misspecification of the values of V_{ic} and V_T is discussed below.

The choice of V_{ic} and V_T does not seem to affect our conclusions substantially. For example, if V_{ic} is changed to V'_{ic} with V_T kept fixed, then numerical results show that $\alpha - \beta$ remains unchanged, νX_u changes to

$$\nu X'_u \approx \frac{V_{ic}}{V'_{ic}} \nu X_u,$$

and $1 - \beta/\alpha$ changes to

$$\left(1 - \frac{\beta}{\alpha}\right)' \approx \frac{V'_{ic}}{V_{ic}}\left(1 - \frac{\beta}{\alpha}\right).$$

(See [10] for similar results.) Similarly, if V_T is changed to V'_T but V_{ic} is kept fixed, then $\alpha - \beta$ and νX_u remain unchanged and $1 - \beta/\alpha$ changes to

$$\left(1 - \frac{\beta}{\alpha}\right)' \approx \frac{V_T}{V'_T}\left(1 - \frac{\beta}{\alpha}\right),$$

or α changes to $\alpha' \approx (V'_T / V_T)\alpha$.

The maximum likelihood estimates of the parameters and the corresponding 95% confidence intervals (based on the observed information matrix) are given in Table 1. The 95% confidence intervals based on profile likelihood and computed using the algorithm of Venzon and Moolgavkar [18] are similar. The estimates obtained by using the original analysis [10], which assumed that the foci were spherical, are also presented in Table 1 for the sake of comparison. The estimates of $\alpha - \beta$ range from 0.0089 per cell per day in the control group to 0.0237

TABLE 1

Maximum Likelihood Estimates with 95% Confidence Intervals

	Arbitrary shape assumption		Spherical shape assumption	
Parameter	Estimate	Confidence interval	Estimate	Confidence interval
c_1	3.9125	(2.541, 5.978)	12.010	(8.571, 16.580)
c_2	3.2484	(2.323, 4.525)	18.908	(14.860, 23.750)
b_1	0.00890	(0.00792, 0.01000)	0.00723	(0.00645, 0.00810)
b_2	0.00037	(0.00033, 0.00041)	0.00026	(0.00023, 0.00029)
β/α	0.9114	(0.8577, 0.9461)	0.9965	(0.9950, 0.9975)
σ^2	0.2563	(0.1784, 0.3682)		

cell^{-1} day^{-1} in the highest dose (40 ppm) group. These values correspond to the estimated cell division rates of 0.1 and 0.267 cell $^{-1}$ day $^{-1}$, respectively. The background cell division rates, as derived from direct measurements of labeling indices (see [19]), are approximately 0.025 cell^{-1} day $^{-1}$. Thus, the estimates in this analysis are biologically more plausible than those (2.057 and 5.029) obtained by a reanalysis of the same data using the method of Moolgavkar et al. [10]. The estimates presented in that paper held β/α fixed at 0.99, which was not done in the reanalysis. Under the assumption that a typical cell (normal or initiated) has radius 12 μm, a cubic centimeter of the tissue contains approximately 1.38×10^{8} cells. As the estimates of νX_0 range from 3.9125 to 133.8485 day $^{-1}$ cm^{-3}, the corresponding estimates of ν range from 2.835×10^{-8} cell^{-1} day^{-1} in the control group of 9.699×10^{-7} cell^{-1} day^{-1} at 40 ppm. All these estimates, however, should be accepted with some caution as they depend on assumptions about the volume of a cell (normal or initiated) and, in this example, also about the volume of the tissue blocks.

The likelihood function was numerically maximized using the Davidon–Fletcher–Powell algorithm [20]. In order to calculate (6) and (7) for evaluation of the likelihood function for different parameter values, we used numerical integration. We used a combination of Legendre–Gauss and Laguerre–Gauss integration schemes (see [20]) for this purpose over a range beyond which V_f has negligible probability (since all of them are very small compared to 1). The idea is to come up with a suitable integration scheme that is stable enough with regard to minor changes and covers almost the whole range of relevant volume fraction (V_f) values. To calculate the truncation probability (6), the Legendre–Gauss scheme was first used over a range with a lower bound corresponding to the volume fraction of one initiated cell ($V_f = V_{ic}/V_T$) and the upper bound corresponding to that of 1000 initiated cells, and then the Laguerre–Gauss scheme was used with the last Gauss point located at $V_f = 0.001$. Results were insensitive to the latter choice. For the calculation of the density of an area fraction a_f [to be precise, the numerator of (7)], we used a similar combination of Legendre–Gauss and Laguerre–Gauss schemes except that the middle point was a chosen as $a_f^{3/2}$, which is an approximation to the point where the corresponding integrand, as a function of v_f, has a maximum. The stability of this numerical integration was monitored by computing the integral of $g(v_f)$, given by (9), at each iteration of likelihood maximization. Stable results were obtained with 40 Gauss points for each integration involved.

To assess the fit of the model, we compare, in Figure 1, the empirical (observed) distribution function $\hat{F}(x)$ based on all the observed area

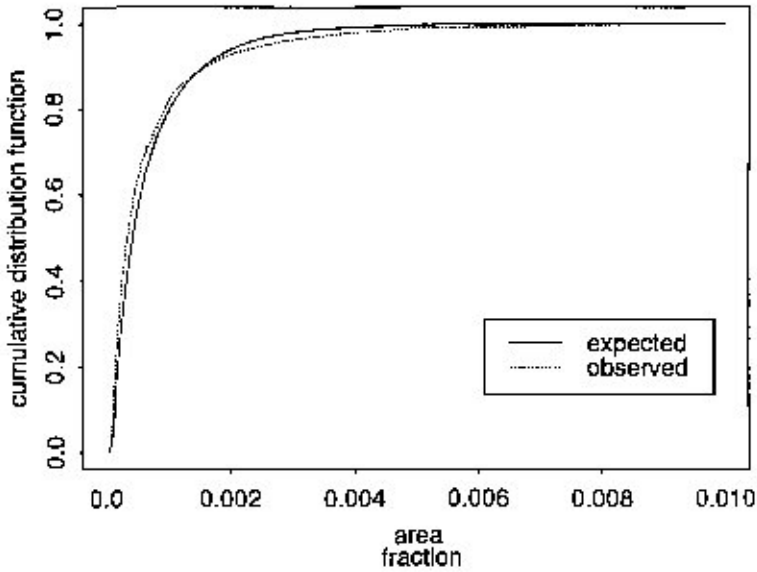

FIG. 1. Empirical (observed) and model-based (expected) cumulative distribution function of area fraction.

fractions with the model-based (expected) distribution function $F_0(x)$ of area fraction calculated as

$$F_0(x) = \frac{\sum_{\text{all obs.}} n_2(i) \times P_i\{\text{area fraction of a focus} \leqslant x\}}{\sum_{\text{all obs.}} n_2(i)},$$

$$= \frac{1}{\sum_{\text{all obs.}} n_2(i)} \sum_{\text{all obs.}} n_2(i) \frac{H(x) - H(\delta/A_i)}{H(\Delta/A_i) - H(\delta/A_i)},$$

where $n_2(i)$ denotes the value of N_2, the number of transections, A_i that of A, area of the cross section, and P_i the probability distribution under the threshold and truncation constraints at the time and dose for the ith animal. Figure 1 gives evidence that the model describes the data well.

DISCUSSION

As we remarked in the previous section, the threshold and truncation constraints do not appear to influence greatly the parameter estimates. If these constraints can be ignored, it is immediately possible to use a least squares method not requiring the distributional assumption (4), as

discussed earlier. It should be noted that when observations on only the number of foci are available, one still needs assumption (4) to use result (11). However, without the threshold and truncation constraints, we use $\Lambda(t)E[V_f^{1/3}]$ for the Poisson parameter instead of $\eta_2(t)$ [see (11)], and hence (4) is not required. Sometimes only the total area fraction of all the foci taken together may be reported. The truncation constraint on the area of a single focus does not extend to the total area of all the foci, making it difficult to analyze such data. Without the threshold consideration, this problem does not arise. One can then analyze such data using results similar to (3), (4), and (5), where v_f is to be treated as the total volume fraction of all the clones taken together. The distribution of the corresponding random variable V_f can then be obtained from that of the total number of initiated cells, W say, which in the case of constant parameters, can be proved to have a negative binomial distribution with

$$p_W(w) = \binom{K+w-1}{K-1}\left(\frac{P}{Q}\right)^w\left(1-\frac{P}{Q}\right)^K, \qquad w = 0,1,2,\ldots,$$

where $P = \theta/(1-\theta)$, $Q = 1+P$, and $K = -\Lambda(t)/[\log(1-\theta)]$. One may want to use a least squares type of method as well, without requiring (4) and using (1) and (2) only, in which case one can check that α and β are not estimable but only $\alpha - \beta$ is.

Note that assumption (4) is central to the whole analysis. We believe that this is a reasonable assumption in the absence of any knowledge regarding the shapes of the clones. However, the estimates are not very different when (4) is replaced by a uniform distribution or a beta distribution with the same mean and variance. Normality in (4) is a reasonable choice having an edge over the uniform distribution in terms of plausibility and over the beta distribution in terms of computational difficulty.

We have not considered the thickness (T_h, say) of the cross sections in our analysis, assuming it to be negligible. Clearly, this will affect the probability of the cross section intersecting the clone. As in [14], the probability of intersection given $V_f = v_f$ can be assumed to be $[v_f^{1/3} - T_h/V_T^{1/3}]/[1 - T_h/V_T^{1/3}]$.

Although the analysis presented in this paper gives more plausible estimates of cell division rates than an analysis based on the assumption of spherical foci, the estimated division rates still appear to be too high, as remarked in the Example section. We should emphasize, however, that our assumptions regarding cell volume (V_{ic}) and volume of tissue block (V_T) certainly affect our estimates. For the analysis in this paper, we assumed that the radius of a single altered cell is 12 μm because

this was the assumption that was made in the original analysis by Moolgavkar et al. [10]. If, however, the radius of an altered cell is taken to be 14 μm as reported in [21], then the estimates of cell division rates are substantially decreased. For example, the division rate in the foci of control animals is estimated to be 0.063 cell^{-1} day^{-1} as opposed to 0.1 cell^{-1} day^{-1}. Also, our estimate of the volume of the tissue block (V_T) is likely to be an overestimate because of the way the sections were cut, which also results in overestimation of the cell division rate, as noted in the preceding section. As discussed there, information on the volumes of the tissue blocks and of a single altered cell in addition to the areas of cross sections is required for a proper analysis using the method in this paper. Experimentalists should be encouraged to record this information.

Finally, the model for cell alteration and clonal growth used in this paper is perhaps the simplest biological model for these processes. The assumptions made by this model, such as homogeneity of division and death rates within foci and the independence of these rates of the size of foci, require further investigation.

This work has been supported by NCI grant 47658 and by an American Cancer Society International Cancer Research Fellowship from UICC to Dr. Dewanji.

REFERENCES

1 H. W. Kunz, H. A. Tennekes, R. E. Port, M. Schwarz, D. Lorke, and G. Schaude, Quantitative aspects of chemical carcinogenesis and tumor promotion in liver, *Environ. Health Perspect.* 50:113–122 (1983).

2 M. Schwarz, D. Pearson, A. Buchmann, and W. Kunz, The use of enzyme-altered foci for risk assessment of hepatocarcinogenesis, in *Biologically Based Methods for Cancer Risk Assessment* (NATO ASI Series A: Life Sci., Vol. 159), C. C. Travis, Ed., Plenum, New York, 1989, pp. 31–39.

3 H. A. Campbell, Y. D. Xu, M. H. Hanigan, and H. C. Pitot, Application of quantitative stereology to the evaluation of phenotypically heterogeneous enzyme-altered foci in the rate liver, *J. Nat. Cancer Inst.* 76:751–767 (1986).

4 D. Nychka, G. Wahba, S. Goldfarb, and T. Pugh, Cross validated spline methods for the estimation of three-dimensional tumor size distributions from observations on two-dimensional cross-sections, *J. Am. Stat. Assoc.* 79:832–846 (1984).

5 A. Dewanji, D. J. Venzon, and S. H. Moolgavkar, A stochastic two-stage model for cancer risk assessment. II. The number and size of premalignant clones, *Risk Anal.* 9:179–187 (1989).

6 E. G. Luebeck and S. H. Moolgavkar, Stochastic analysis of intermediate lesions in carcinogenesis experiments, *Risk Anal.* 11:149–157 (1991).

7 S. H. Moolgavkar and D. J. Venzon, Two-event models for carcinogenesis: incidence curves for childhood and adults tumors, *Math. Biosci.* 47:55–77 (1979).

8 S. H. Moolgavkar and A. G. Knudson, Mutation and cancer: a model for human carcinogenesis, *J. Nat. Cancer Inst.* 66:1037–1052 (1981).
9 S. H. Moolgavkar and E. G. Luebeck, Two-event models for carcinogenesis: biological, mathematical and statistical considerations, *Risk Anal.* 10:323–341 (1990).
10 S. H. Moolgavkar, E. G. Luebeck, M. de Gunst, R. E. Port, and M. Schwarz, Quantitative analysis of enzyme-altered foci in rat hepatocarcinogenesis experiments. I. Single agent regimen, *Carcinogenesis* 11:1271–1278 (1990).
11 E. G. Luebeck, S. H. Moolgavkar, A. Buchmann, and M. Schwarz, Effects of polychlorinated biphenyls in rat liver: quantitative analysis of enzyme altered foci, *Toxicol. Appl. Pharmacol.* 111:469–484 (1991).
12 D. S. Wicksell, The corpuscle problem, Part I, *Biometrika* 17:87–97 (1925).
13 D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic Geometry and Its Applications*, Wiley, New York, 1987.
14 N. Keiding and L. Anderson, Estimation of the size distribution of fibrillar centres in nucleoli—an example of the "Swiss cheese" problem in stereology, *Biometrics* 48:449–458 (1992).
15 A. Baddeley, A limit theorem for statistics of spatial data, *Adv. Appl. Probab.* 12:447–461 (1980).
16 S. Mase, Asymptotic properties of stereological estimators of volume fraction for stationary random sets, *J. Appl. Probab.* 19:111–126 (1982).
17 A. Dewanji, S. H. Moolgavkar, and E. G. Luebeck, Two mutation model for carcinogenesis: joint analysis of premalignant and malignant lesions, *Math. Biosci.* 104:97–109 (1991).
18 D. J. Venzon and S. H. Moolgavkar, A method for computing profile-likelihood based confidence intervals, *Appl. Stat.* 37:87–94 (1988).
19 S. H. Moolgavkar and E. G. Luebeck, Interpretation of labeling indices in the presence of cell death, *Carcinogenesis* 13:1007–1010 (1992).
20 W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes: The Art of Scientific Computing*, Cambridge Univ. Press, New York, 1986.
21 E. M. Jack, P. Bentley, F. Bieri, S. F. Muakkassah-Kelly, W. Stäubli, J. Suter, F. Waechter, and L. M. Cruz-Orive, Increase in hepatocyte and nuclear volume and decrease in the population of binucleated cells in preneoplastic foci of rat liver: a sterological study using the nucleator method, *Hepatology* 11:286–297 (1990).